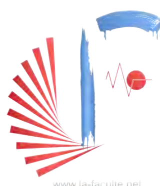
In the name of Allah, the Most Gracious, the Most Merciful






Devoir sur feuille n° 4.

Exercice 1 :

Pour étudier l'importance de la chenille processionnaire du pin d'Alep, qui constitue la principale espèce du barrage vert, on a compté le nombre de pins parasités dans 100 lots de 8 arbres chacun. Soit X le nombre de pins parasités par lot.

1) Quelle est la loi de probabilité de X ?

2) Si la moyenne est de 2 arbres parasités par lot, déterminer la probabilité p d'avoir un arbre parasité.

3) Déterminer la probabilité de ne trouver aucun arbre parasité par lot.

4) Quelle est la probabilité de trouver au moins un arbre parasité par lot ?

5) Quelle est la probabilité de trouver moins de 2 arbres non parasités par lot ?

Exercice 2 :

Pour tester l'efficacité de 5 médicaments, on a compté pour 100 individus, le nombre de réactions favorables à chacun des traitements. On a obtenu :

Réactions favorables	0	1	2	3	4	5
Effectifs	4	11	22	35	20	8

Soit X le nombre de réactions favorables par individu.

1) Déterminer la moyenne et l'écart type de la variable X.

2) On suppose que X suit une loi binomiale.

a- Donner ses paramètres n et p.

b- Calculer les probabilités d'avoir 2 réactions favorables, et au plus 2 réactions favorables.

Exercice 3 :

5% des pièces d'une fabrique ont un défaut. On prélève 80 pièces. Soit X la VAD qui compte le nombre de pièces défectueuses.

1) Donner la loi de probabilité de X.

2) En utilisant une approximation convenable, calculer la probabilité d'avoir :

a- 2 pièces défectueuses, b- au moins 3 pièces défectueuses.

CORRECTION DU DEVOIR SUR FEUILLE N°4.

Exercice 1:

1/ soit $\mathcal{E}$: "examiner 1 pin" → S : "Il est parasité" $P(S) = p = ?$; → E : "Il est sain" $P(E) = 1-p = ?$

- S et E sont exclusifs ($S \cap E = \phi$) et complémentaires ($P(S)+P(E)=1$)
- Examiner 8 pins correspond à 8 répétitions indépendantes de $\mathcal{E}$

X = nombre de pins parasités par lot de 8 pins = 0, 1, ..., 8.

$X \rightsquigarrow \mathcal{B}(n,p) \iff P(X=k) = C_8^k \, p^k (1-p)^{8-k}$, $k = 0, 1, \dots, 8$.
$n = 8$; $p = ?$

2/ $X \rightsquigarrow \mathcal{B}(n,p) \Rightarrow E(X) = np = 2 \Rightarrow p = \frac{2}{n} = \frac{2}{8} = \frac{1}{4}$
donc $P(X=k) = C_8^k \left(\frac{1}{4}\right)^k \left(\frac{3}{4}\right)^{8-k}$.

3/ $P(X=0) = (3/4)^8 = 0{,}1001$
4/ $P(X \geq 1) = 1 - P(X=0) = 1 - 0{,}1001 = 0{,}8999$
5/ soit A : "Trouver moins de 2 arbres non parasités"
= 0 ou 1 arbre non parasité
= 8 ou 7 arbres parasités.

$\Rightarrow P(A) = P(X=7) + P(X=8) = C_8^7 (1/4)^7 (3/4) + C_8^8 (1/4)^8$
$= 0{,}000366 + 0{,}000015 = 0{,}000381$.

Exercice 2:

1/

x	0	1	2	3	4	5	Σ
n_i	4	11	22	35	20	8	100
$n_i x_i$	0	11	44	105	80	40	280
$n_i x_i^2$	0	11	88	315	320	200	934

$\bar{x} = \frac{1}{n} \sum n_i x_i = \frac{280}{100} = 2{,}8$

$\sigma^2_{ech} = \frac{1}{n} \sum n_i x_i^2 - \bar{x}^2 = \frac{934}{100} - (2{,}8)^2 = 1{,}5 \Rightarrow \sigma_{ech} = 1{,}225$.

2/ X = nombre de réactions favorables = 0, 1, ..., 5
a/ $X \rightsquigarrow \mathcal{B}(n,p)$ avec $n = 5$ et $p = ?$

$E(X) = np = \bar{x} \Rightarrow p = \frac{\bar{x}}{n} = \frac{2{,}8}{5} = 0{,}56$

b/ $P(X=k) = C_5^k (0{,}56)^k (0{,}44)^{5-k}$, $k=0,1,\dots,5$

i/ $P(X=2) = C_5^2 (0{,}56)^2 (0{,}44)^3 = 0{,}26714$

ii/ $P(X \leq 2) = P(X=0) + P(X=1) + P(X=2)$

avec $P(X=0) = (0{,}44)^5 = 0{,}01649$

$P(X=1) = 5 \times 0{,}56 \times (0{,}44)^4 = 0{,}10495$

$\Rightarrow P(X \leq 2) = 0{,}38858$.

Exercice 3 :

1/ soit $\mathcal{E}$: " prélever une pièce. " $\rightarrow$ S : " défectueuse. " $P(S) = 0{,}05$; $\rightarrow$ E : " saine. " $p(E) = 0{,}95$

. S et E sont exclusifs ($S \cap E = \phi$) et complémentaires ($P(S) + P(E) = 1$)

. Prélever 80 pièces = 80 répétitions indépendantes de $\mathcal{E}$.

X = nombre de pièces défectueuses $= 0, 1, \dots, 80$

$X \sim \mathcal{B}(80 ; 0{,}05) \Leftrightarrow P(X=k) = C_{80}^k (0{,}05)^k (0{,}95)^{80-k}$

2/ $\left.\begin{array}{l} X \sim \mathcal{B}(n,p) \\ n = 80 \geq 30 \\ p = 0{,}05 \leq 0{,}1 \\ np = 4 \leq 18 \end{array}\right\} \Rightarrow X \sim \mathcal{P}(\lambda)$, $\lambda = np = 4$

$\Downarrow$

$P(X=k) = e^{-4} \cdot \frac{4^k}{k!}$

a/ $P(X=2) = e^{-4} \cdot \frac{4^2}{2!} = 0{,}146525$

b/ $P(X \geq 3) = 1 - P(X<3) = 1 - [P(X=0) + P(X=1) + P(X=2)]$

$= 1 - [e^{-4} + 4e^{-4} + P(X=2)]$

$= 1 - [0{,}091578 + 0{,}146525]$

$= 0{,}761897$.

Devoir sur feuille n° 5.

Exercice 1 :(Les questions 2), 3) et 4) sont indépendantes).

Dans une population donnée, la taille Y des individus suit une loi normale d'espérance μ et de variance σ^2 inconnues.

1) On choisit un individu au hasard. Quelle est la probabilité que sa taille soit supérieure à $\mu + 0.44\sigma$?

2) On considère un échantillon de 20 individus.

a- Quelle est la loi de la variable X = nombre d'individus de taille supérieure à $\mu+0.44\sigma$?

b- Quelle est la probabilité d'avoir :

i) Plus d'un individu dont la taille est supérieure à $\mu + 0.44\sigma$?

ii) Au plus 17 individus de taille inférieure ou égale à $\mu + 0.44\sigma$?

3) Si on considère un échantillon de 100 individus, quelle est la probabilité d'avoir :

a- Au moins 25 individus de taille supérieure à $\mu + 0.44\sigma$?

b- Au plus 80 individus de taille inférieure ou égale à $\mu + 0.44\sigma$?

4) On désire, maintenant, déterminer les valeurs de μ et σ. On sait que :

- 84.13% des individus ont une taille supérieure à 155 cm, et que
- 2.275% des individus ont une taille inférieure à 140 cm.

a- Trouver alors les valeurs de μ et de σ.

b- Déterminer N tel que 80% des tailles soient comprises entre 140 cm et N cm (N>140).

Exercice 2 :

La glycémie des sujets d'une population est supposée distribuée suivant une loi normale de paramètres $\mu = 1$ g. et $\sigma = 0.2$ g.

1) Quelle est pour un individu, la probabilité d'avoir une glycémie comprise entre 0.7g et 1.2g ?

2) Trouver le seuil x_0 pour lequel 60% de la population a une glycémie qui lui est supérieure.

3) Trouver 2 valeurs a et b, symétriques par rapport à la moyenne μ, telles que la glycémie soit comprise entre ces 2 valeurs avec une probabilité égale à 0.9.

CORRECTION du C.F. N°3.

Exercice 1 :

Y = taille des individus $\quad Y \rightsquigarrow \mathcal{N}(\mu;\sigma) \iff U = \frac{Y-\mu}{\sigma} \rightsquigarrow \mathcal{N}(0;1)$

1/ $P(Y > \mu + 0{,}44\sigma) = 1 - P(Y \leq \mu + 0{,}44\sigma) = 1 - P\left(U \leq \frac{\mu + 0{,}44\sigma - \mu}{\sigma}\right)$

$= 1 - \Phi(0{,}44) = 1 - 0{,}67 = 0{,}33.$

2/ a/ soit $\mathcal{E}$ = taille d'1 individu $\begin{cases} E_1 : \text{"taille} > \mu + 0{,}44\sigma\text{"} & P(E_1) = 0{,}33 \\ E_2 : \text{"taille} \leq \mu + 0{,}44\sigma\text{"} & P(E_2) = 0{,}67 \end{cases}$

E_1 et E_2 sont exclusifs ($E_1 \cap E_2 = \emptyset$) et complémentaires ($P(E_1) + P(E_2) = 1$)

Taille de 20 individus = 20 répétitions indépendantes de $\mathcal{E}$.

X = nombre d'individus de taille $> \mu + 0{,}44\sigma = 0, 1, 2, \ldots, 20$

alors $X \rightsquigarrow \mathcal{B}(20; 0{,}33) \iff P(X=k) = C_{20}^{k}(0{,}33)^k(0{,}67)^{20-k}$, $k = 0, 1, \ldots, 20$

b/ i°/ $P(X > 1) = 1 - P(X \leq 1) = 1 - P(X=0) - P(X=1) = 1 - 0{,}67^{20} - 20 \times 0{,}33 \times 0{,}67^{19} = 1 - 0{,}0036 = 0{,}$

ii°/ soit A : "Au plus 17 individus de taille $\leq \mu + 0{,}44\sigma$"

donc A : "Au moins (20-17) individus de taille $> \mu + 0{,}44\sigma$" $= \{X \geq 3\}$

$\Rightarrow P(A) = P(X \geq 3) = 1 - P(X < 3) = 1 - P(X \leq 1) - P(X = 2) = 0{,}9964 - 0{,}0153 = 0{,}9811.$

3/ soit Z = nombre d'individus de taille $> \mu + 0{,}44\sigma$, parmi 100.

alors $Z \rightsquigarrow \mathcal{B}(100; 0{,}33)$ (même raisonnement que pour X)

a/ $P(Z \geq 25) = P(Z = 25) + \cdots + P(Z = 100)$ Trop long, voire impossible.

Mais, de plus $p = 0{,}33 \in]0{,}2; 0{,}8[$ et $n = 100 \geq 30 \Rightarrow Z \rightsquigarrow \mathcal{N}(m; s)$

où $m = n.p = 100 \times 0{,}33 = 33$ et $s = \sqrt{np(1-p)} = 4{,}7$

D'où $P(Z \geq 25) = 1 - \Phi\left(\frac{25-33}{4{,}7}\right) = 1 - \Phi(-1{,}7) = \Phi(1{,}7) = 0{,}9554.$

b/ soit B : "Au plus 80 individus de taille $\leq \mu + 0{,}44\sigma$"

B : "Au moins (100-80) individus de taille $> \mu + 0{,}44\sigma$" $= \{Z \geq 20\}$

donc $P(B) = P(Z \geq 20) = 1 - \Phi\left(\frac{20-33}{4{,}7}\right) = 1 - \Phi(-2{,}77) = \Phi(2{,}77) = 0{,}9972.$

4/ On doit résoudre le système : $\begin{cases} P(Y > 155) = 0{,}8413 \\ P(Y < 140) = 0{,}02275 \end{cases}$

a/ . $P(Y > 155) = 1 - \Phi\left(\frac{155-\mu}{\sigma}\right) = 1 - \Phi(u_1) = 0{,}8413 = \Phi(-u_1)$ avec $u_1 = \frac{155-\mu}{\sigma}$

donc $\Phi(-u_1) = 0{,}8413 = \Phi(1) \Rightarrow -u_1 = 1 \Rightarrow u_1 = \frac{155-\mu}{\sigma} = -1 \Rightarrow \underline{\mu = 155 + \sigma}$ (1)

.. $P(Y < 140) = \Phi\left(\frac{140-\mu}{\sigma}\right) = \Phi(u_2) = 0{,}02275$ avec $u_2 = \frac{140-\mu}{\sigma}$

donc $\Phi(u_2) = 0{,}02275 = 1 - \Phi(-u_2) \Rightarrow \Phi(-u_2) = 1 - 0{,}02275 = 0{,}97725 \simeq \Phi(2)$

$\Rightarrow -u_2 = 2 \Rightarrow u_2 = \frac{140-\mu}{\sigma} = -2 \Rightarrow \underline{\mu = 2\sigma + 140}$ (2)

Les relations ① et ② donnent :

$\mu = \sigma + 155 = 2\sigma + 140 \Rightarrow \underline{\sigma = 15 cm}$ et $\underline{\mu} = 15 + 155 = \underline{170 cm}$.

b/ $M = ?$ tel que $P(140 \leq X \leq M) = 0,8$

$\Leftrightarrow P\left(\frac{140-170}{15} \leq U \leq \frac{M-170}{15}\right) = \Phi\left(\frac{M-170}{15}\right) - \Phi(-2) = 0,8$

$\Rightarrow \Phi\left(\frac{M-170}{15}\right) = 0,8 + 1 - \Phi(2) = 0,8228$

on a alors : $[\Phi(0,92) =] 0,8212 < [\Phi\left(\frac{M-\mu}{\sigma}\right) =] 0,8228 < 0,8238 [= \Phi(0,93)]$

$\Rightarrow 0,92 < \frac{M-\mu}{\sigma} < 0,93$

$\Rightarrow \frac{M-\mu}{\sigma} = 0,93 \Rightarrow M = \mu + 0,93\sigma = 183,95$ cm. $\underline{M = 183,95 cm}$

Exercice 2 :

Soit X = glycémie. $X \sim \mathcal{N}(1 ; 0,2) \Leftrightarrow U = \frac{X-1}{0,2} \sim \mathcal{N}(0 ; 1)$

$m = 1 ; \sigma = 0,2$

1/ $P(0,7 \leq X \leq 1,2) = \Phi\left(\frac{1,2-1}{0,2}\right) - \Phi\left(\frac{0,7-1}{0,2}\right) = \Phi(1) - \Phi(-1,5)$

$= \Phi(1) - 1 + \Phi(1,5) = 0,8413 - 1 + 0,9332 = 0,7745$.

2/ $x_0 = ?$ tel que $P(X \geq x_0) = 0,6 = 1 - P(X < x_0) = 1 - \Phi\left(\frac{x_0 - m}{\sigma}\right) = \Phi\left(\frac{m - x_0}{\sigma}\right)$

$\Rightarrow \Phi\left(\frac{m-x_0}{\sigma}\right) = 0,6 \simeq \Phi(0,25)$ car : $0,5987 < 0,6 < 0,6026$

$\Leftrightarrow \Phi(0,25) < \Phi\left(\frac{m-x_0}{\sigma}\right) < \Phi(0,26)$

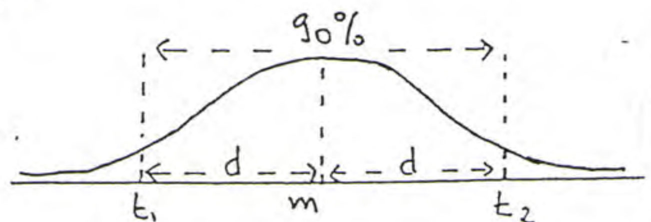
$\Rightarrow 0,25 < \frac{m-x_0}{\sigma} < 0,26$

$\Rightarrow \frac{m-x_0}{\sigma} = 0,25 \Rightarrow x_0 = m - 0,25\sigma$

$\Rightarrow x_0 = 1 - 0,25 \times 0,2 = 0,95$ g

3/ on doit résoudre $P(t_1 < X < t_2) = 0,9$

avec $\begin{cases} t_1 = m - d \\ t_2 = m + d \end{cases}$



donc $P(m - d < X < m + d) = 0,9$

i.e : $P\left(-\frac{d}{\sigma} < U < \frac{d}{\sigma}\right) = 0,9$

i.e : $\Phi\left(\frac{d}{\sigma}\right) - \Phi\left(-\frac{d}{\sigma}\right) = 2\Phi\left(\frac{d}{\sigma}\right) - 1 = 0,9$

$\Rightarrow \Phi\left(\frac{d}{\sigma}\right) = \frac{1+0,9}{2} = 0,95 \simeq \Phi(1,65)$

$\Rightarrow \frac{d}{\sigma} = 1,65 \Rightarrow d = 1,65\sigma = 1,65 \times 0,2 = 0,33 \Rightarrow \begin{cases} t_1 = m - d = 0,67 \text{ g} \\ t_2 = m + d = 1,33 \text{ g}. \end{cases}$

Devoir sur feuille n° 7.

Exercice 1 :

On étudie le poids des nouveau-nés durant les 12 premiers mois de leur naissance.

1) Sur un échantillon dont l'âge ne dépasse pas 6 mois, on observe une moyenne m_1 et un écart type $\sigma_1 = 5$ kg. La moyenne de la population de cette catégorie d'âge a été estimée par l'intervalle de confiance [6.10 ; 8.90], au taux de sécurité de 95%.
Déterminer la moyenne m_1 et la taille n_1 de cet échantillon.

2) On observe un autre échantillon de 45 bébés dont l'âge est supérieur à 6 mois. On a obtenu une moyenne $m_2 = 9.5$ kg et une variance $\sigma_2^2 = 36$ (kg)².
En comparant les intervalles de confiance, peut-on dire que le poids des nouveau-nés augmente significativement à partir du 6ème mois ? (Sinon, faire le test approprié).

3) Déterminer la taille minimale n des 2 échantillons ($n_1 = n_2 = n$) qu'il faut choisir pour pouvoir affirmer que cette augmentation est significative au risque $\alpha = 5\%$, par :

a- la méthode de l'écart réduit,

b- la méthode de recouvrement (intersection) des intervalles de confiance.

(On supposera que les moyennes et variances observées reste inchangées).

Exercice 2 :

Dans une enquête sur l'étiologie du cancer broncho-pulmonaire, on a interrogé des sujets atteints de ce cancer et des témoins non malades sur leur consommation de tabac :

- Sur 100 cancéreux, 80 d'entre eux sont des fumeurs,
- Sur 400 témoins, il y a 100 non fumeurs.

1) En comparant les intervalles de confiance, peut-on affirmer au niveau de confiance de 95% que les 2 proportions de non fumeurs ne diffèrent pas significativement ? (Faire le test correspondant si nécessaire).

2) En supposant que les 2 tailles d'échantillon sont égales, et qu'on retrouve les mêmes proportion déjà observées, quelle serait la taille minimale nécessaire pour une différence significative au risque de 5% ? Procéder par la méthode :

a- de l'écart réduit,

b- des intervalles de confiance.

Quelle méthode est-il plus raisonnable de choisir ? Justifier.

Correction du DF n°7.

Exercice 1.

1/ $IC(M_1)_{95\%} = [m_1 - h_1 ; m_1 + h_1] = [6,10 ; 8,90] \Rightarrow m_1 = \frac{6,10+8,90}{2} = 7,5$ kg.

$h_1 = \frac{8,90-6,10}{2} = 1,4$ et $h_1 = t_\alpha \frac{\sigma_1}{\sqrt{n_1 - 1}} \Rightarrow \sqrt{n_1-1} = \frac{t_\alpha \sigma_1}{h_1}$

$\Rightarrow n_1 - 1 = \left(\frac{t_\alpha \sigma_1}{h_1}\right)^2 \Rightarrow n_1 = \left(\frac{t_\alpha \sigma_1}{h_1}\right)^2 + 1 = \left(\frac{1,96 . 5}{1,4}\right)^2 + 1 = 50$

2/ $n_2 = 45 ; m_2 = 9,5 ; \sigma_2 = 6 \Rightarrow IC(M_2)_{95\%} = [m_2 - h_2 ; m_2 + h_2]$, $h_2 = \frac{1,96 . 6}{\sqrt{44}} = 1,77$
$= [7,73 ; 11,27]$

$IC(M_1) \cap IC(M_2) = [7,73 ; 8,90] \neq \phi$
de plus $m_1 = 7,5 \notin IC(M_2)$. donc on ne peut pas conclure directement.
Test de l'écart-réduit :
- Ho : Il n'y a pas de ≠ce significative entre les 2 moyennes observées.
- $\varepsilon = \frac{|m_1 - m_2|}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} = \frac{9,5 - 7,5}{\sqrt{\frac{25}{50} + \frac{36}{45}}} = 1,75 < 1,96$ donc Ho retenue à 95%

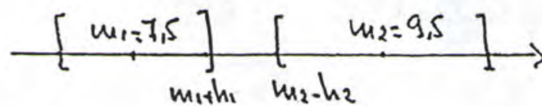
Donc on ne peut pas conclure, au t.s. de 95%, que le poids des bébés augmente significativement à partir du 6e mois.

3/ L'augmentation sera significative si : a/ $\varepsilon \geq t_\alpha$
b/ $IC(M_1) \cap IC(M_2) = \phi$

a/ $\varepsilon = \frac{|m_1 - m_2|}{\sqrt{\frac{\sigma_1^2 + \sigma_2^2}{n_1}}} = \sqrt{n} . \frac{|m_1 - m_2|}{\sqrt{\sigma_1^2 + \sigma_2^2}} \geq t_\alpha$

$\Leftrightarrow \sqrt{n} \geq \frac{t_\alpha \sqrt{\sigma_1^2 + \sigma_2^2}}{|m_1 - m_2|} \Rightarrow n \geq \left(\frac{t_\alpha}{m_1 - m_2}\right)^2 (\sigma_1^2 + \sigma_2^2) = 58,58$
$\Rightarrow n_{min} = 59$

b/ $IC(M_1) \cap IC(M_2) = \phi$



$\Rightarrow m_1 + h_1 < m_2 - h_2$
$\Rightarrow h_1 + h_2 < m_2 - m_1$
$\Rightarrow \frac{t_\alpha \sigma_1}{\sqrt{n-1}} + \frac{t_\alpha \sigma_2}{\sqrt{n-1}} < m_2 - m_1 \Rightarrow \frac{t_\alpha}{\sqrt{n-1}} (\sigma_1 + \sigma_2) < m_2 - m_1$
$\Rightarrow \sqrt{n-1} > \frac{t_\alpha (\sigma_1 + \sigma_2)}{m_2 - m_1} \Rightarrow n > \left(\frac{t_\alpha (\sigma_1 + \sigma_2)}{m_2 - m_1}\right)^2 + 1 = 117,21$
$\Rightarrow n_{min} = 118$

Exercice 2 :

Ech1 : $n_1 = 100$ cancéreux, 80 fumeurs $\Rightarrow k_1 = 20$ non-fumeurs
$\Rightarrow P_1 = \frac{k_1}{n_1} = \frac{20}{100} = 0,2$

Ech2 : $n_2 = 400$ témoins, $k_2 = 100$ non-fumeurs $\Rightarrow P_2 = \frac{k_2}{n_2} = \frac{100}{400} = 0,25$

1/ $IC(P_{01})_{95\%} = [P_1 - h_1 ; P_1 + h_1] = [0,1216 ; 0,2784]$ car $h_1 = t_\alpha \sqrt{\frac{P_1(1-P_1)}{n_1}} = 0,0784$

$IC(P_{02})_{95\%} = [P_2 - h_2 ; P_2 + h_2] = [0,2076 ; 0,2924]$ car $h_2 = 1,96 \sqrt{\frac{0,25 . 0,75}{400}} = 0,0424$

$IC(P_{01}) \cap IC(P_{02}) = [0,2076 ; 0,2784] \neq \phi$
de plus $P_1 = 0,2 \notin IC(P_2)$
donc on ne peut pas conclure directement.

Test de l'écart-réduit :
H_0 : Il n'y a pas de $\neq$ce significative entre les 2 %tages.
soit $p = \frac{k_1 + k_2}{n_1 + n_2} = \frac{120}{500} = 0,24$

$$\varepsilon = \frac{|P_1 - P_2|}{\sqrt{\frac{p(1-p)}{n_1} + \frac{p(1-p)}{n_2}}} = \frac{0,05}{\sqrt{\frac{0,24 \times 0,76}{100} + \frac{0,24 \times 0,76}{400}}} = 1,05 < 1,96$$

on peut donc affirmer au s.s. de 95% que les %tage de non-fumeurs ne diffèrent pas significativement.

2/ Différence significative $\Rightarrow$ a/ $\varepsilon \geq t_\alpha$
b/ $IC(P_{01}) \cap IC(P_{02}) = \phi$.

a/ $$\varepsilon = \frac{|P_1 - P_2|}{\sqrt{\frac{p(1-p)}{n_0} + \frac{p(1-p)}{n_0}}} = \frac{|P_1 - P_2|}{\sqrt{\frac{2p(1-p)}{n_0}}} = \frac{\sqrt{n_0} . |P_1 - P_2|}{\sqrt{2p(1-p)}}$$ avec $p = \frac{nP_1 + nP_2}{n+n} = \frac{P_1 + P_2}{2} = 0,225$

donc $\varepsilon \geq t_\alpha \Leftrightarrow \frac{\sqrt{n_0} . |P_1 - P_2|}{\sqrt{2p(1-p)}} \geq t_\alpha \Leftrightarrow \sqrt{n_0} \geq \frac{t_\alpha \sqrt{2p(1-p)}}{|P_1 - P_2|}$

donc $n_0 \geq \left(\frac{t_\alpha}{P_1 - P_2}\right)^2 . 2p(1-p) = 535,9 \Rightarrow n_{0min} = 536$.

b/ $IC(P_{01}) \cap IC(P_{02}) = \phi \Rightarrow P_1 + h_1 < P_2 - h_2 \Rightarrow h_1 + h_2 < P_2 - P_1$

$\Rightarrow t_\alpha \sqrt{\frac{P_1(1-P_1)}{n_0}} + t_\alpha \sqrt{\frac{P_2(1-P_2)}{n_0}} = \frac{t_\alpha}{\sqrt{n_0}} \left(\sqrt{P_1(1-P_1)} + \sqrt{P_2(1-P_2)}\right) < P_2 - P_1$

$\Rightarrow \sqrt{n_0} > \frac{t_\alpha}{P_2 - P_1} . \left(\sqrt{P_1(1-P_1)} + \sqrt{P_2(1-P_2)}\right) \Rightarrow n_0 > \left[\frac{t_\alpha}{P_2 - P_1} . \left(\sqrt{P_1(1-P_1)} + \sqrt{P_2(1-P_2)}\right)\right]^2$

donc $n_0 > 1066,29 \Rightarrow n_{0min} = 1067$.
La méthode de l'écart réduit est plus précise. De plus elle nécessite moins d'individus (536 au lieu de 1067).

Devoir sur feuille n° 8.

Exercice 1 :

On administre à des cobayes 4 injections intradermiques de teinture de goudron, à des intervalles de temps assez espacés, et on observe le nombre de réactions allergiques pour chaque cobaye. On obtient le tableau suivant :

Nombre de réactions	0	1	2	3	4
Nombre de cobayes	52	84	42	16	6

1. a- Calculer le nombre total d'injections administrées.
b- Calculer le nombre total de réactions allergiques.
c- En déduire la fréquence expérimentale de réactions allergiques.

2. Soit X le nombre de réactions allergiques observées chez un cobaye. Peut-on ajuster la distribution observée à une loi binomiale ? (Prendre $\alpha = 0.05$)

3. a- Quel est le nombre moyen $\bar{x}$ de réactions allergiques ?
b- Si X a pour écart type $\sigma = 1.5$, Peut-on considérer que l'écart entre $\bar{x}$ et 1 est significatif à 95% ? (Ici on a pris $\mu = 1$).
c- Quelle taille d'échantillon minimale n_0 doit-on prendre pour pouvoir affirmer que cet écart est significatif à 99% ?

4. Soit P_0 la probabilité d'avoir moins de 2 réactions allergiques, et P le pourcentage observé d'individus ayant moins de 2 réactions allergiques. Peut-on affirmer, à 99%, que P et P_0 diffèrent significativement ? Procéder par 2 méthodes.

Exercice 2 :

On a dosé la quantité de créatine (en mg/100cm^3 d'urine) chez 80 hommes. On a obtenu les résultats suivants :

Quantité de créatine	[2.5 ; 3.5[	[3.5 ; 4.5[	[4.5 ; 5.5[	[5.5 ; 6.5[	[6.5 ; 7.5[
Effectifs	3	19	30	22	6

1. Déterminer la moyenne m et la variance σ^2.

2. Peut-on admettre, au niveau de confiance de 95%, que la distribution ci-dessus obéit à la loi normale N(m, σ) ?

CORRIGE DF 8

Exercice I :

1- a/ nombre total d'injections administrées : $N = 200 \times 4 = 800$

b/ " " de réactions allergiques : $n_1 = \sum n_i x_i = 240$

c/ soit P la fréquence expérimentale de réactions allergiques :

alors $P = \frac{\text{nombre de cas favorables}}{\text{nombre de cas possibles}} = \frac{n_1}{N} = \frac{240}{800} = 0{,}3$

2- X = nombre de réactions allergiques $= 0, 1, 2, 3, 4$.

$X \sim \mathcal{B}(4\,;0{,}3) \Rightarrow P_i = P(X=i) = C_4^i (0{,}3)^i (0{,}7)^{4-i}$.

test d'ajustement :

soit : H_0 = "Il y a ajustement de la distribution observée à la loi $\mathcal{B}(4;0{,}3)$."

calcul des P_i :

$P_0 = P(X=0) = C_4^0 (0{,}3)^0 (0{,}7)^4 = (0{,}7)^4 = 0{,}2401 \Rightarrow C_0 = 200 P_0 = 48{,}02$

$P_1 = P(X=1) = C_4^1 (0{,}3)^1 (0{,}7)^3 = 4 \times 0{,}3 \times (0{,}7)^3 = 0{,}4116 \Rightarrow C_1 = 82{,}32$

$P_2 = P(X=2) = C_4^2 (0{,}3)^2 (0{,}7)^2 = 0{,}2646 \Rightarrow C_2 = 52{,}92$

$P_3 = P(X=3) = 4 \times (0{,}3)^3 . 0{,}7 = 0{,}0756 \Rightarrow C_3 = 15{,}12$

$P_4 = P(X=4) = C_4^4 . (0{,}3)^4 = (0{,}3)^4 = 0{,}0081 \Rightarrow C_4 = 1{,}62 < 5$

on remarque que $C_4 = 1{,}62 < 5$, donc on regroupe les 2 dernières classes.

X	0	1	2	3 ou 4	Σ
P_i	0,2401	0,4116	0,2646	0,0756 0,0081	1
O_i	52	84	42	16 6 → 22	200
$C_i = 200 P_i$	48,02	82,32	52,92	15,12 1,62 → 16,74	200
$\frac{(O_i - C_i)^2}{C_i}$	0,3299	0,0343	2,2533	1,6528	$\chi^2 = \sum_{i=1}^{4} \frac{(O_i - C_i)^2}{C_i} = 4{,}2703$

p est estimé $\Rightarrow \ell = 1 \Rightarrow ddl = k - 1 - \ell = 4 - 1 - 1 = 2$

$\alpha = 0{,}05 \Rightarrow \chi^2_{5\%, 2ddl} = 5{,}991 > \chi^2 \Rightarrow H_0$ retenue à 95%

3- a/ nombre moyen de réactions allergiques : $\bar{x} = \frac{1}{n} \sum_{i=1}^{5} n_i x_i = \frac{240}{200} = 1{,}2$

b/ $\mu = 1$, $\sigma = 1{,}5$

comparaison d'une moyenne observée $(\bar{x})$ à une moyenne théorique (μ)

soit H_0 : "Il n'y a pas de $\neq$ce significative entre les moyennes observée et théorique"

$\varepsilon = \frac{|\bar{x} - \mu|}{\sigma / \sqrt{n}} = \frac{(1{,}2 - 1)}{1{,}5} . \sqrt{200} = 1{,}886 < t_\alpha = 1{,}96 \Rightarrow H_0$ retenue à 95%

Cf $n_0 = ?$ tq $\varepsilon \geq t_\alpha = 2{,}58$ ($\alpha = 1\%$)

$$\varepsilon = \frac{|\bar{x}-\mu|}{\sigma/\sqrt{n_0}} = \sqrt{n_0}\cdot\frac{|\bar{x}-\mu|}{\sigma} \geq t_\alpha \Rightarrow \sqrt{n_0} \geq \frac{t_\alpha \cdot \sigma}{|\bar{x}-\mu|}$$

$$\Rightarrow n_0 \geq \left(\frac{t_\alpha \sigma}{\bar{x}-\mu}\right)^2 = \left(\frac{2{,}58 \times 1{,}5}{0{,}2}\right)^2 = 374{,}42 \Rightarrow n_{0\,min} = 375$$

4-
- $P_0 = P(x < 2) = P(x=0) + P(x=1) = 0{,}2401 + 0{,}4116$ (voir tableau) $= 0{,}6517$

- $P = \frac{O_0 + O_1}{n} = \frac{52+84}{200} = \frac{136}{200} = 0{,}68$.

- Comparaison d'1 %tage observé à 1 %tage théorique

$n\,p_0 = 200 \times 0{,}6517 = 130{,}34 \geq 5$
$n(1-p_0) = 200(1-0{,}6517) = 69{,}66 \geq 5$
donc les 2 tests sont applicables.

1ère Méthode : test de l'écart-réduit :

Ho : Il n'y a pas de ≠ce significative entre les 2 %tages observé et théorique

$$\varepsilon = \frac{|P-P_0|}{\sqrt{\frac{P_0(1-P_0)}{n}}} = \frac{|0{,}6517 - 0{,}68|}{\sqrt{\frac{0{,}6517 \times 0{,}3483}{200}}} = 0{,}84 < t_\alpha = 2{,}58 \Rightarrow \text{Ho acceptée}$$

2è Méthode : test du χ^2 (de conformité)

Ho : Il y a conformité entre les résultats observés et théoriques.

	Moins de 2 réactions	2,3,0,4 réactions	Σ
P_i	0,6517	0,3483	1
O_i	136	64	200
$C_i = 200P_i$	130,34	69,66	200
$\frac{(O_i-C_i)^2}{C_i}$	0,246	0,46	$\chi^2 = 0{,}706$

$$\chi^2 = \sum_{i=1}^{2} \frac{(O_i - C_i)^2}{C_i} = 0{,}706$$

$ddl = 2-1 = 1$, $\alpha = 0{,}01$ $\Rightarrow \chi^2_{1\%;1ddl} = 6{,}635 > \chi^2$

donc Ho acceptée au risque $\alpha = 1\%$

EXERCICE II :

$n_1 = 100$ cancéreux, 80 fumeurs $\Rightarrow k_1 = 20$ non fumeurs $\Rightarrow P_{01} = \frac{20}{100} = 0{,}2$

$n_2 = 400$ témoins, $k_2 = 100$ non fumeurs $\Rightarrow P_{02} = \frac{100}{400} = 0{,}25$

comparaison de 2 %tages observés.

1ère Méthode : test du χ^2 d'indépendance.

Ho : Il y a indépendance entre la consommation de tabac et le cancer des bronches.

	Fumeurs	Non F	n_i
cancéreux	80 / 76	20 / 24	100
témoins	300 / 304	100 / 96	400
n'_j	380	120	500

$\frac{(O_{ij}-C_{ij})^2}{C_{ij}}$	F	$\bar{F}$	Σ
C	0,211	0,667	0,876
T	0,053	0,167	0,22
			1,096

$C_{ij} = \frac{n_i \cdot n'_j}{N} \geq 5$ vérifié

$$\chi^2 = \sum_{i=1}^{2}\sum_{j=1}^{2} \frac{(O_{ij}-C_{ij})^2}{C_{ij}} = 1{,}096$$

$ddl = (2-1)(2-1) = 1$, $\alpha = 5\% \Rightarrow \chi^2_{5\%;1ddl} = 3{,}841 > \chi^2$

$\Rightarrow$ Ho acceptée à 95%.

2è Méthode : test de l'écart-réduit.

Ho : Il n'y a pas de ≠ce significative entre les 2 %tages observés de non fumeurs.

$$\varepsilon = \frac{|P_{01} - P_{02}|}{\sqrt{\frac{P(1-P)}{n_1} + \frac{P(1-P)}{n_2}}} = \frac{0{,}25 - 0{,}2}{\sqrt{0{,}24 \times 0{,}76\left(\frac{1}{100} + \frac{1}{400}\right)}} = 1{,}047$$

avec $P = \frac{k_1 + k_2}{n_1 + n_2} = \frac{120}{500} = 0{,}24$

$\varepsilon = 1{,}047 < t_\alpha = 1{,}96 \Rightarrow$ Ho acceptée à 95%.

Exercices : Tests de Conformité, d'Ajustement.

Exercice 1 :

Le taux moyen de prothrombine (substance contenue dans le sang, qui participe à sa coagulation) dans une population normale est de 20 mg/100ml de plasma. Dans un échantillon de 50 patients atteints d'une déficience en vitamine K, le taux moyen de prothrombine est de 18 mg/100ml, l'écart type observé est de 4 mg/100ml.

1) Cet échantillon est-il représentatif de la population, au t.s. de 0.95 ? Répondre en utilisant 2 méthodes.

2) Que répondez-vous au risque de 0.01 ?

3) Si on ne dispose que de 10 patients, quelle serait votre conclusion ? Quelle hypothèse supplémentaire doit-on alors admettre ?

Exercice 2 :

Dans un échantillon de 50 individus on a calculé la moyenne de cholestérol, m = 160 cg. Sachant que la moyenne de cholestérol dans la population est de 156 cg et son écart type de 14 cg, à partir de quel taux de sécurité (1- α) peut-on affirmer que l'échantillon est représentatif de la population ?

Exercice 3 :

On estime à 80% la réussite d'un traitement T contre une maladie virale. Au cours de l'année, 120 malades se sont présentés à un dispensaire et on reçu ce traitement. On constate que le traitement a été efficace pour 84 sujets. On veut savoir si l'efficacité du traitement annoncée à 80% est vérifiée.

1) Quels tests peut-on utiliser ?

2) Quelle est l'hypothèse nulle à tester ?

3) Après calcul, quel est le jugement d'interprétation ?

Exercice 4 :

Dans un échantillon de 10000 oiseaux, on trouve 5240 mâles.

1) Peut-on admettre, au risque de 5%, l'hypothèse de 2 proportions égales de mâles et de femelles dans la population d'origine ? Procéder par 2 méthodes :

a- le test de l'écart réduit, **b-** le test du Khi -deux.

2) Quelle est votre réponse au niveau de confiance de 99% ?

3) Déterminer la taille maximale n_0 qui permettrait d'accepter, au risque $\alpha = 0.01$, l'hypothèse considérée.

Exercice 5 :(Synthèse 2004)

Dans une population masculine adulte, on a déterminé les proportions suivantes :

Non fumeurs : 50% Fumeurs de cigarettes : 40%

Fumeurs de pipe : 3% Fumeurs de cigarettes et de pipe : 7%

Sur un échantillon de 200 sujets de sexe masculin atteints d'une maladie déterminée, on a observé : 120 non fumeurs, 5 fumeurs de pipe, 70 fumeurs de cigarettes et 5 fumeurs de pipe et cigarettes. Cette répartition diffère-t-elle de celle obtenue sur la population, au risque de 5% ?

Exercice 6 :

Après de nombreuses années d'étude clinique, on a constaté que pour les malades atteints d'un cancer aplasique broncho-pulmonaire primitif, la survie sans traitement, une fois le diagnostic posé, se distribue de la manière suivante :

Survie en mois	< 6	6 à 12	12 à 24	> 24
Fréquence de survie	0.45	0.35	0.15	0.05

Pour 60 malades soumis à un traitement T, associant une poly-chimiothérapie première suivie d'une radiothérapie, on a observé les résultats suivants :

Survie en mois	< 6	6 à 12	12 à 24	> 24
Nombre de malades	6	24	12	18

Au vu de ces résultats, au risque de 5%, le traitement T est-il efficace ?

Exercice 7 :

Une variété de souris présente des tumeurs spontanées avec un taux P = 20%. Sur 20 souris traitées on observe 8 tumeurs spontanées. On se demande si la proportion observée diffère significativement de P, au t.s. de 99%.

Exercice 8 :

Un laboratoire reçoit par groupes de 4, des souris grises et blanches, destinées aux expérimentations. Il reçoit 200 groupes de 4 souris et observe les résultats suivants :

Nombre de souris blanches	0	1	2	3	4
Nombre de groupes	15	65	70	35	15

1) Une hypothèse génétique simple conduit à supposer que les couleurs grise et blanche sont également réparties entre les souris.
Soit X le nombre de souris blanches dans un groupe de 4 souris.

a- Vérifier que X suit une loi binomiale.
b- Déterminer les probabilités d'observer les différentes valeurs de X.
c- En déduire les effectifs théoriques correspondant à chaque valeur de X.

2) Au niveau de confiance de 95%, la répartition observée dans le tableau précédent donne-t-elle un résultat compatible avec la loi de X ?

Exercice 9 :

Le tableau suivant représente, pour une période de 100 jours, le nombre de jours durant lesquels se sont produits des accidents automobiles.

Nombre d'accidents	0	1	2	3	4	5
Nombre de jours	42	36	14	6	2	0

1) a- Calculer la moyenne observée m.
b- Peut-on ajuster la distribution observée à une loi de Poisson de paramètre m, à 99% ?
2) Peut-on l'ajuster à la loi de Poisson de paramètre $\lambda = 1.2$? ($\alpha = 0.05$)

Exercice 10 :

La pesée de 50 nouveau-nés dans une maternité a permis d'établir le tableau suivant :

Classes	2.0 – 2.4	2.4 – 2.8	2.8 – 3.2	3.2 – 3.6	3.6 – 4.0
Effectifs	6	11	19	10	4

1) Déterminer la moyenne m et la variance σ^2 correspondant à ces bébés.
2) Peut-on considérer, au risque de 5%, que la distribution ci-dessus obéit à la loi normale $N(m\ ;\sigma)$?
Et au risque de 1% ?

En effet :

$\mu \in IC(\mu)_{1-\alpha} = [\bar{x} - h ; \bar{x} + h] \iff \bar{x} - h \leq \mu \leq \bar{x} + h$

$-t_\alpha \frac{\sigma_{ech}}{\sqrt{n-1}} = -h \leq \mu - \bar{x} \leq h = t_\alpha \frac{\sigma_{ech}}{\sqrt{n-1}}$

$-t_\alpha \leq \frac{\mu - \bar{x}}{\sigma_{ech}/\sqrt{n-1}} \leq t_\alpha$

$\Rightarrow \varepsilon = \frac{|\mu - \bar{x}|}{\sigma_{ech}/\sqrt{n-1}} \leq t_\alpha$

- Exo 1 -

Pop : $\mu = 20$

Ech : $n = 10$ $\bar{x} = 18$ $\sigma_{ech} = 4$.

① - on compare μ à $\bar{x}$

a/ Test de l'écart réduit

H_0 : l'ech est représentatif à pop.

$\varepsilon = \frac{|\bar{x} - \mu|}{\sigma_{ech}/\sqrt{n-1}} = \frac{\sqrt{49}}{4} \cdot |20 - 18| = 3,5 \geq t_{5\%} = 1,96 \Rightarrow H_0$ rejetée à 95%

$\Rightarrow$ l'éch n'est pas représentatif de la pop à 95%.

b. $\mu \notin IC(\mu)_{95\%}$

$IC(\mu)_{95\%} = [\bar{x} - h ; \bar{x} + h]$ $h = t_\alpha = \frac{\sigma_{ech}}{\sqrt{n-1}} = \frac{1,96 \times 4}{7} = 1,12$

$= [16,88 ; 19,12]$

$\mu = 20 \notin [16,88 ; 19,12]$

② - $\alpha = 0,01 \Rightarrow t_\alpha = 2,58 \Rightarrow \varepsilon = 3,5 > 2,58 \Rightarrow$ m réponse à 99%.

③ - $n = 10$. σ_{ech} inconnu $\Rightarrow \varepsilon = \frac{|\bar{x} - \mu|}{\sigma_{ech}/\sqrt{n-1}} = \frac{2}{4} \times \sqrt{9} = 1,5$

⇓

on doit admettre que la variable étudié est normale.

$ddl = n - 1 = 9 \Rightarrow t^*_{5\%} = 2{,}262$ ②

II) Comparaison d'1 proportion obs à 1 proportion théo ($n \geq 30$)

$nP_0 \geq 5$

$n(1 - P_0) \geq 5$

Pop $\rightarrow P_0$ inconnue

Ech $\rightarrow n, K \Rightarrow P = \frac{K}{n}$

Q : l'éch et -issu de la pop

R : Test de l'écart réduit

H_0 : Il n'y a pas de ≠ce entre les 2 pop obs

$$\varepsilon = \frac{|P - P_0|}{\sqrt{\frac{P_0(1-P_0)}{n}}}$$

-Exo 3-

Pop : $P_0 = 0{,}8 = 80\%$

éch : $n = 120$, $K = 84 \Rightarrow P = \frac{K}{n} = \frac{84}{120} = 0{,}7$

① - Test d'écart réduit : comparaison d'1 proportion obs et 1 proportion théo :

② - H_0 : Il n'y a pas de ≠ significative entre les 2 proportions.

③ - $\varepsilon = \frac{|P - P_0|}{\sqrt{\frac{P_0(1-P_0)}{n}}} = \frac{|0{,}8 - 0{,}7|}{\sqrt{0{,}7 \frac{(1-0{,}7)}{120}}} = 2{,}74 > t_{5\%} = 1{,}95 \Rightarrow H_0$ rejetée à 5%.

$\Rightarrow$ l'efficacité du traitement à 80% n'est pas vérifiée.

-Ex02-

Pop : $\mu = 156$ $\sigma = 14$

ech : $n = 50$ $m = 160$

l'ech est représentatif de la population si $\varepsilon < t_\alpha$

$$\varepsilon = \frac{|\mu - \bar{x}|}{\sigma/\sqrt{n}} = \frac{|160-156|}{14/\sqrt{50}} = 2{,}02 < t_\alpha$$

$\varepsilon < t_\alpha \Rightarrow \Phi(\varepsilon) < \Phi(t_\alpha) = 1 - \frac{\alpha}{2} \Rightarrow \alpha < 2(1 - \Phi(\varepsilon))$

$\alpha < 2(1 - \Phi(2{,}02))$

$\alpha < 2(1 - 0{,}97831)$

$\alpha < 0{,}0434$

$\alpha < 0{,}0434 \Rightarrow -\alpha > -0{,}0434 \Rightarrow 1-\alpha > 1 - 0{,}0434 \Rightarrow 1-\alpha > 0{,}9566$

-Ex04-

	Mâles	Femelles
Pop :	$P_0 = \frac{1}{2}$	$1 - P_0 = 1/2$
Ech :	5240	4760

$\Rightarrow P_0 = \frac{5240}{10000} = 0{,}524$

① a- Test de l'écart-réduit : (comparaison d'1 %age)

H_0 = il n'y a pas de ≠ce significative entre les 2 proportions.

$$\varepsilon = \frac{|P - P_0|}{\sqrt{P_0(1-P_0)/n}} = \frac{|0{,}524 - 0{,}5|}{\sqrt{\frac{(0{,}5)^2}{10^4}}} = \frac{0{,}024}{0{,}5} \times 10^2 = 2 \times 0{,}024 \times 100 = 4{,}8 > t_{5\%} = 1{,}96$$

$\Rightarrow H_0$ rejetée à 95%

③

- Exo 5 - H_0 : il ya conformité entre les résultats obs et théo

Pop :

	$\bar{F}$	FC	FP	FCP	Σ
P_i	0,5	0,4	0,03	0,07	1
O_i	120	70	5	5	200
C_i	100	80	6	14	n=200 C_i tous ≥ 5 vérifié
$\frac{(O_i-C_i)^2}{C_i}$	4	1,25	0,17	5,78	$\chi^2 = \sum_{i=1}^{4} \frac{(O_i-C_i)^2}{C_i} = 11,21$

ddl = K − 1 = 4 − 1 = 3 ⇒ $\chi^2_{3ddl} = 7,815$

$\chi^2 = 11,21 > \chi^2_\alpha$ ⇒ H_0 rejetée à 95% ⇒ la répartition obs ≠ de celle obtenue sur la pop.

- Exo 6 -

H_0 : Il ya conformité ....

	<6	6 à 12	12 à 24	>24	
P_i	0,45	0,35	0,15	0,05	1
O_i	6	24	12 (30)	18	n=60
$C_i = 60P_i$	27	21	9 (12)	3	$C_4 = 3 < 5$ ⇒ on regroupe les 2 dernières données
$\frac{(O_i+C_i)^2}{C_i}$	16,33	0,43	27.		n=60 $\chi^2 = \sum_{i=1}^{3} \frac{(O_i-C_i)^2}{C_i} = 43,76$

ddl = K − 1 = 3 − 1 = 2 ⇒ $\chi^2_{2ddl} = 5,991$

$\chi^2 >> \chi^2_\alpha$ ⇒ H_0 rejetée à 95%.

⇒ traitement efficace.

-Exo7-

Pop : $P_0 = 0,2$ — Tumeurs — $\overline{\text{Tumeurs}}$: $1 - P_0 = 0,8$

Ech : $n = 20$, $k = 8$; $P = \frac{k}{n} = 0,4$

l'écart réduit

H_0 : Il y a conformité entre les résultats obs et théo.

	T	$\overline{T}$	Σ
P_i	0,2	0,8	1
O_i	8	12	n = 20
$\frac{(O_i + C_i)^2}{C_i}$	4	16	n = 20
$\frac{(\lvert O_i - C_i \rvert - \frac{1}{2})^2}{C_i}$			$\chi^2 = 3,83$

$C_i < 5$ (n'est pas regroupé)

⇒ Correct° de Yates

$\chi^2 = \Sigma \frac{(\lvert O_i - C_i \rvert - 1/2)^2}{C_i} = 3,83$

ddl = 2 - 1 ⇒ $\chi^2_{t} = 6,635$.

b- Test du χ^2 de conformité :

H_0 : Il y a conformité entre les résultats obs et théo

	M	F	Σ
P_i	1/2	1/2	1
O_i	5240	4760	n=10000
$\frac{(O_i - C_i)^2}{C_i}$	11,52	11,52	$C_i \geq 5$ ⟹ vérifié

$\chi^2 = \sum_1^k \frac{(O_i - C_i)^2}{C_i} = 23,04$

$ddl = K-1 = 2-1 = 1 \Rightarrow \chi^2_{1ddl, C} = 3,846 < \chi^2 \Rightarrow H_0$ rejetée à 95%.

R! $\chi^2 = \varepsilon^2$

$\chi^2_{\alpha, ddl} = (t_\alpha)^2$.

$[N(0,1)]^2 = \chi^2_{1ddl}$

$O_i \sim N(C_i, \sqrt{C_i})$

$\frac{(O_i - C_i)^2}{C_i} = \left(\frac{O_i - C_i}{\sqrt{C_i}}\right)^2 \sim (N(0,1))^2 = \chi^2_{ddl}$

$\chi^2_{corr} = \sum_{i=1}^{2} \frac{(|O_i - C_i| - \frac{1}{2})^2}{C_i}$

(2) $\varepsilon = 4,8 > t_C = 2,58 \Rightarrow$ rien ne change à 99%

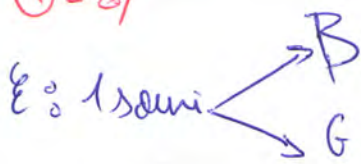
(3) - n_0 ? t.q $\varepsilon < t_\alpha$:

$\Leftrightarrow \varepsilon = \frac{|P - P_0|}{\sqrt{\frac{P_0(1-P_0)}{n_0}}} < t_\alpha \Leftrightarrow \sqrt{n_0}\frac{|P-P_0|}{\sqrt{P_0(1-P_0)}} < t_\alpha \Leftrightarrow \sqrt{n_0} < t_\alpha \cdot \frac{|P-P_0|}{\sqrt{P_0(1-P_0)}}$

$n_0 < \left(\frac{t_\alpha}{P-P_0}\right)^2 (P_0(1-P_0)) \left(\frac{2,58}{0,024}\right)^2 (0,5)^2$

$\Rightarrow$ $n_{0\,max} = 2889$





① - a/

$\mathcal{E}$: 1 souris → B $p(B) = 0,5$; → G $p(G) = 0,5$

B et G sont exclusifs et complémentaires. On fait 4 répétitions indépendantes de $\mathcal{E}$.

X : nbr de souris blanches = 0, 1, 2, 3, 4 $X \sim B(n,p)$ $n = 4$, $p = 0,5$

$$P(X=k) = C_n^k \, p^k (1-p)^{n-k} = C_4^k \left(\frac{1}{2}\right)^4 = \frac{C_4^k}{2^4} = \frac{C_4^k}{16}$$

b/ $P(X=0) = \frac{1}{16} = 0,0625$

$P(X=1) = \frac{C_4^1}{16} = 0,25$

$P(X=2) = \frac{C_4^2}{16} = \frac{6}{16} = 0,375$

$P(X=3) = \frac{C_4^3}{16} = 0,25$

$P(X=4) = \frac{C_4^4}{16} = 0,0625$

c/ $C_i = NP_i = 200 P_i$

$C_0 = 200 \times 0,0625 = 12,5$

$C_1 = 200 \times 0,25 = 50$

$C_2 = 200 \times 0,375 = 75$

$C_3 = 50$

$C_4 = 12,5$.

② - Test d'ajustement à la loi $B(n,p)$

H_0 : il y a ajustement de la loi obs à la loi $B(n,p)$ $n = 4$, $p = \frac{1}{2}$

X	O_i	P_i	$C_i = 200 P_i$	$(O_i - C_i)^2 / C_i$
0	15	0,0625	12,5	0,5
1	65	0,25	50	4,5
2	70	0,375	75	0,33
3	35	0,25	50	2,5
4	15	0,0625	12,5	0,5
Σ	200	1	200	$\chi^2 = \sum_{i=1} \frac{(O_i - C_i)^2}{C_i} = 10,33$

$p = \frac{1}{2}$ donnée ⇒ $l = 0$

⇒ $ddl = K - 1 - l = 5 - 1 - 0 = 4$

$\chi^2_{4 ddl, 5\%} = 9,488$

$\chi^2 > \chi^2_\alpha$ ⇒ H_0 rejetée à 5%.

- Exo 9 -

X	O_i	$O_i X_i$	P_i	$C_i = 100 P_i$
0	42	0	0,4066	40,66
1	36	36	0,3659	36,59
2	14	28	0,1647	16,47
3	6	18	0,0494	4,94
4	2	8	0,0111	1,11
5	0	0	0,0020	0,20
	100	90	0,9997 ≃ 1	99,79 ≃ 100

(classes 3, 4, 5 : 6,25)

$C_3, C_4, C_5 < 5 \Rightarrow$ on regroupe les 3 dernières classes

λ estimé par $m = 0,9 \Rightarrow l = 1$

$\Rightarrow ddl = k_1 - l = 4 - 1 - 1 = 2 \Rightarrow \chi^2_{\alpha, ddl} = 9,21$

$\chi^2 = \sum_{1}^{4} \frac{(O_i - C_i)^2}{C_i} = 0,91$

$\chi^2 < \chi^2_\alpha$

$\Rightarrow H_0$ accepté à 99%

donc y a ajustement à la loi poisson

a/ $M = \frac{1}{n} \sum n_i x_i = \frac{90}{100} = 0,9$

b/ Ajustement à la loi $P(\lambda)$:

H_0 : il y a ajustement de la loi $P(m)$ $m = 0,9$

- Calcul des P_i : $X \sim P(m = 0,9) \Rightarrow P(X = k) = e^{-0,9} \frac{(0,9)^k}{k!}$

$P_0 = (P(X = 0)) = e^{-0,9} \frac{0,9^0}{0!} = e^{-0,9} = 0,4066$

$P_1 = (P(X = 1)) = e^{-0,9} \times 0,9 = \lambda P(X=0) = 0,3659$

$P_2 = P(X = 2) = \frac{\lambda}{2} P(X = 1) = 0,1647$

$P_3 = P(X = 3) = \frac{\lambda}{3} P(X = 2) = 0,0494$

$P_4 = P(X = 4) = \frac{\lambda}{4} P(X = 3) = 0,0111$

$P_5 = P(X = 5) = \frac{\lambda}{5} P(X = 4) = 0,0020$

② -